

CLUSTERING NEGARA BERDASARKAN SKOR PENGENDALIAN KONSUMSI TEMBAKAU MENGGUNAKAN ALGORITMA DBSCAN

JOKO RIYONO¹, CHRISTINA ENI PUJIASTUTI², AINA LATIFA RIYANA PUTRI³

^{1,2}Universitas Trisakti, Jl. Kyai Tapa No.1 Grogol, Jakarta, ³Universitas Diponegoro, Jl.Prof.Sudarto No.13 Tembalang, Semarang
Email : jokoriyono@trisakti.ac.id

ABSTRACT

Smoking is an activity that has a detrimental impact on the health of individuals, families, communities and the environment, both directly and indirectly. Therefore, it is necessary to control global tobacco consumption. The World Health Organization (WHO) in the Framework Convention on Tobacco Control (FTCT) developed the MPOWER strategy: a which countries can use to control tobacco consumption. In this study a clustering analysis of tobacco use control will be carried out based on the MPOWER score: a from WHO for each country using the Density Based Spatial Clustering of Applications with Noise (DBScan) algorithm. DBSCAN is a cluster formation algorithm based on the level of distance density between objects in a dataset. Using a density radius of 1.72 with a minimum point of 4 objects obtained from the kNNdisplot function on Rstudio produces 4 clusters, 75 data as noise, and the Davies Bouldin-Index value as the best cluster validity is 1.08118.

Keywords: *Clustering Analysis, Density Radius, Cluster Validity*

ABSTRAK

Merokok merupakan aktivitas yang mempunyai dampak merugikan bagi kesehatan individu, keluarga, masyarakat dan lingkungan, baik secara langsung maupun tidak langsung. Oleh karena itu diperlukan upaya pengendalian konsumsi tembakau secara global. World Health Organization (WHO) dalam Framework Convention on Tobacco Control (FTCT) menyusun strategi MPOWER: a yang dapat digunakan oleh negara-negara dalam melakukan pengendalian konsumsi tembakau. Pada penelitian ini akan dilakukan analisis clustering pengelompokan pengontrolan penggunaan tembakau berdasarkan skor penilaian MPOWER: a dari WHO untuk setiap negara menggunakan algoritma Density Based Spatial Clustering of Applications with Noise (DBScan). DBSCAN adalah algoritma pembentukan cluster berdasarkan tingkat kerapatan jarak antar objek dalam dataset. Menggunakan radius kerapatan 1.72 dengan minimum points sebesar 4 objek yang didapat dari fungsi kNNdisplot pada Rstudio menghasilkan 4 cluster, 75 data sebagai noise, dan nilai Davies Bouldin-Index sebagai validitas cluster terbaik sebesar 1.08118.

Kata Kunci : Analisis Clustering, Radius Kerapatan, Validitas Cluster

1. PENDAHULUAN

Tanaman tembakau merupakan salah satu tanaman tropis asli Amerika dengan kelompok tumbuhan dari genus *Nicotina* yang daunnya digunakan sebagai elemen utama bahan baku pembuatan rokok. Selain tembakau, sebatang rokok mengandung lebih dari 4000 jenis bahan kimia beracun yang menimbulkan risiko serius bagi kesehatan manusia [1] seperti Nikotin yang dapat menyebabkan kecanduan dan masalah kesehatan termasuk penyakit jantung, penyakit paru-paru, dan kanker [2]. Karena memiliki zat adiktif yang menyebabkan adiksi (ketagihan) dan depedensi (ketergantungan) bagi orang yang menghisapnya, dalam sebuah artikel milik Badan Narkotika Nasional (BNN) [3] rokok termasuk golongan NAPZA (Narkotika, Psikotropika, Alkohol, dan Zat Adiktif). Selain itu, Tar dalam rokok pada penelitian [4] menunjukkan memiliki kaitan dengan risiko kanker. Jadi, kebiasaan merokok dapat menyebabkan berbagai penyakit seperti kanker, penyakit jantung, penyakit paru paru dan bahkan bisa menyebabkan kematian. Menurut Global Burden of Disease Collaborative Network khususnya di Indonesia pada tahun 2019 terdapat 246.359 jumlah kematian akibat rokok dan pada tahun 2020 terdapat prevalensi orang dewasa merokok yaang berumur diatas 15 tahun sebesar 37,60%. Asap yang dihasilkan dari rokok juga sumber polusi udara yang secara langsung dapat mengganggu kondisi fisik terutama pada anak-anak dan wanita hamil seperti iritasi mata dan hidung, sakit kepala, tenggorokan serak, batuk, kepala pusing dan gangguan pernafasan.

Karena merokok merupakan aktivitas yang berdampak merugikan bagi kesehatan individu, keluarga, masyarakat dan lingkungan, baik secara langsung maupun tidak langsung, diperlukan upaya

pengendalian konsumsi tembakau secara global. World Health Organization (WHO) dalam Framework Convention Tobacco Control (FTCT) menyusun strategi MPOWER: a yang digunakan dalam melakukan upaya pengendalian konsumsi tembakau di beberapa anggotanya seperti adanya peraturan perundang-undangan yang dibuat tiap negara terkait bebas rokok pada beberapa tempat (lingkungan sekolah, restoran, rumah sakit), terkait bentuk promosi atau sponsor iklan tembakau, terkait informasi pajak dan harga penjualan produk tembakau, hingga adanya kampanye anti tembakau [5].

Kementerian Kesehatan Indonesia berdasarkan hasil survei global menyebutkan bahwasannya penggunaan tembakau pada usia dewasa (Global Adult Tobacco Survey-GATS), terjadi peningkatan signifikan jumlah perokok dewasa sebanyak 8,8 juta yaitu 60,3 juta di 2011 menjadi 69,1 juta perokok pada di 2021. Melihat jumlah perokok dan jumlah kematian akibat rokok terus menerus naik khususnya di Indonesia, Pada penelitian ini akan dilakukan analisis clustering pengendalian konsumsi tembakau berdasarkan skor penilaian MPOWER: a dari WHO untuk setiap negara menggunakan algoritma Density Based Spatial Clustering of Applications with Noise (DBSCAN) dan melakukan uji validitas cluster menggunakan metode Davies-Bouldin Index berbantuan software Rstudio. Dengan melakukan analisis kuantitatif MPOWER: a di setiap negara, diharapkan dapat menumbuhkan rasa persaingan dalam menekan prevalensi merokok dan konsumsi tembakau.

DBSCAN melakukan clustering yang memperhatikan konektivitas dari densitas atau algoritma yang masuk dalam kategori density-based clustering, yaitu proses pembentukan cluster dilakukan

berdasarkan tingkat kedekatan/kerapatan jarak antar objek dalam dataset tersebut. Algoritma ini unggul dengan kemampuannya dalam mendeteksi titik-titik outlier atau noise yang ada namun cluster yang terbentuk tergantung dengan nilai input yang diberikan dan mampu menciptakan beberapa cluster yang berbentuk bebas serta acak (tidak bulat) [6]. Clustering menggunakan DBSCAN telah banyak digunakan di beberapa penelitian sebelumnya, seperti pada [7] untuk mengidentifikasi anomali penerbangan dan menunjukkan deteksi anomali tersebut sensitif terhadap bandara. Lalu pada [8] melakukan pengelompokan Cervical Auscultation sebagai metode skrining untuk kelainan menelan dengan membandingkan Algoritma DBSCAN dan Algoritma k-Means. Diperoleh Algoritma DBSCAN memiliki sensitivitas yang lebih tinggi dan dapat mengelompokkan kelainan menelan secara tepat.

Untuk evaluasi clustering, pada penelitian ini dilakukan agar dapat mengetahui seberapa baik kualitas dari hasil clustering. Kevaliditasan hasil cluster dimana kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari cluster yang diikuti dilihat dari Davies-Bouldin Index [9]. Evaluasi dengan Davies-Bouldin Index memiliki skema evaluasi dari internal cluster [10], dimana baik atau tidaknya hasil cluster dilihat dari kuantitas dan kedekatan antar data hasil cluster.

2. METODOLOGI PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif, dengan studi literasi adapun data dikumpulkan menggunakan alat ukur kemudian dianalisis dengan statistik dan kuantitatif. Berikut landasan teori dan metodologi penelitian yang digunakan.

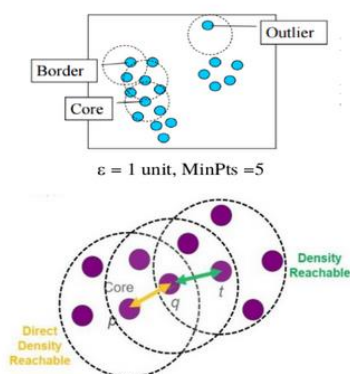
2.1 Algoritma Density-based Spatial Clustering

Algoritma DBSCAN (*Density – based Spatial Clustering of Application with Noise*) merupakan algoritma pengelompokan berdasarkan kepadatan dari lokasi data yang diamati dengan prinsip pengelompokan data yang relatif berdekatan, pemikiran Ester dkk pertama kali diterbitkan pada tahun 1996 [11]. DBSCAN biasanya diterapkan pada data yang banyak noise, bahkan DBSCAN tidak akan mengimpor cluster data apa pun yang dianggap noise [12]. Kategori noise pada algoritma DBSCAN adalah poin-poin dengan kepadatan objek yang sedikit. Jadi prinsip dari algoritma DBSCAN yaitu suatu poin dapat dikelompokkan pada suatu kategori apabila poin tersebut berada dekat dengan poin-poin pada kategori tersebut.

DBSCAN memerlukan dua parameter input sebelum melakukan proses clustering, yaitu epsilon (ϵ) dan skor minimum (MinPts). Epsilon adalah jarak maksimum yang diperbolehkan antara dua data dalam satu cluster dan titik min adalah jumlah minimum data dalam jarak epsilon untuk membentuk sebuah cluster. Metode jarak yang digunakan pada DBSCAN adalah Euclidean distance. Selain epsilon dan MinPts, ada beberapa istilah lain dalam metode DBSCAN, yaitu:

- a *Directly density - reachable*: Pengamatan q berhubungan langsung dengan p, jika p adalah core point dan q merupakan tetangga dari p dalam jangkauan epsilon.
- b *Density-reachable*: Pengamatan q dan x dalam satu cluster namun x bukan

- tetangga dari q dalam jangkauan epsilon.
- c Core point: Merupakan poin inti dimana pengamatan yang memiliki jumlah tetangga lebih atau sama dengan $MinPts$ pada jangkauan epsilon.
- d Border point: Merupakan poin tepi dimana pengamatan memiliki tetangga lebih sedikit atau $MinPts$ namun ia merupakan tetangga atau berada dekat dengan core point.
- e Outlier/noise point: Merupakan poin pencilan dimana pengamatan yang bukan border points maupun core points.



Gambar 1. Terminologi dalam Algoritma DBSCAN

Kelebihan menggunakan metode DBSCAN untuk analisis pengelompokan adalah tidak memerlukan jumlah cluster untuk diberikan sebelumnya, tidak memerlukan asumsi apapun mengenai kepadatan atau varian dalam cluster, dan dapat mendeteksi kelompok dari data yang berjumlah besar serta mengandung noise dan outlier [13].

Selama clustering menggunakan DBSCAN, data akan dikelompokkan dengan cluster tetangga. Sepasang pengamatan dikatakan bertetangga apabila jarak antara dua amatan tersebut kurang dari atau sama dengan epsilon. Secara sederhana cara kerja DBSCAN adalah sebagai berikut:

- a Tentukan parameter $MinPts$ dan epsilon (ϵ).
- b Tandai semua titik data dalam dataset sebagai *unvisited* (belum dikunjungi).
- c Pilih titik P *unvisited* (belum dikunjungi) secara acak dan tandai sebagai *visited* (telah dikunjungi).
- d Hitung jarak antara titik “ p ” terhadap setiap titik lain menggunakan Euclidian distance.

$$d(x, y) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

Dengan $d(x, y)$:ukuran ketidakmiripan; (p_1, p_2) :koordinat titik pusat pada dimensi rendah; (q_1, q_2) :koordinat titik pusat pada dimensi rendah.

- e Jika setidaknya ada sebanyak $MinPts$ titik di dalam radius ketetanggaan P maka cluster baru terbentuk. Jika tidak, P ditandai sebagai noise.
- f Ulangi langkah 4 dan 5 pada setiap titik dalam ketetanggaan sebagai titik P .
- g Ulangi langkah 3 sampai 6 hingga semua amatan diproses.

Evaluasi clustering dilakukan untuk mengetahui seberapa baik kualitas dari hasil clustering. Evaluasi hasil clustering pada penelitian ini menggunakan Davies-Bouldin Index. Metode ini diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979 [14]. Evaluasi dengan Davies-Bouldin Index memiliki skema evaluasi dari internal cluster yang mengukur evaluasi cluster pada suatu metode pengelompokan yang didasarkan pada nilai kohesi dan separasi. Dalam suatu pengelompokan kohesi didefinisikan sebagai nilai kedekatan data terhadap pusat dari clusternya. Sedangkan separasi didefinisikan sebagai seberapa berbeda keterpisahan sebuah cluster dari cluster lain.

Sum of square within cluster (SSW) adalah rumus yang digunakan untuk mendapatkan keterikatan matrik kohesi

dalam sebuah cluster ke-i yang dirumuskan sebagai berikut:

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (2)$$

Keterangan:

m_i = jumlah data dalam cluster ke-i

c_i = centroid cluster ke-i

x_j = data ke-j dalam cluster

Sum of square between cluster (SSB) merupakan persamaan yang digunakan untuk mengetahui separasi antar cluster yang dijumlahkan menggunakan persamaan:

$$SSB_{i,j} = d(c_i, c_j) \quad (3)$$

Keterangan:

c_i = centroid cluster ke-i

c_j = centroid cluster ke-j

$d(c_i, c_j)$ = jarak antara centroid cluster ke-i dengan centroid cluster ke-j

Berdasarkan nilai SSW dan SSB yang didapat, kemudian dilakukan perhitungan rasio (R_{ij}) untuk mengetahui nilai rasio antara cluster ke-i dan cluster ke-j. Cluster yang baik adalah cluster yang memiliki angka kohesi sekecil mungkin dan separasi yang sebesar mungkin. Nilai rasio dihitung menggunakan persamaan sebagai berikut

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (4)$$

Keterangan:

SSW_i = sum of square within cluster pada centroid cluster ke-i

SSW_j = sum of square within cluster pada centroid cluster ke-j

$SSB_{i,j}$ = sum of square between cluster data ke-i dengan j pada cluster yang berbeda.

Nilai rasio yang diperoleh tersebut digunakan untuk mencari nilai Davies Bouldin Index (DBI) dari persamaan berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (5)$$

Keterangan:

k = jumlah cluster yang ditentukan

$R_{i,j}$ = ratio dari nilai SSW dan SSB

Semakin kecil nilai DBI yang diperoleh (non-negatif ≥ 0), maka semakin baik cluster yang diperoleh [15] dari pengelompokan algoritma DBSCAN yang digunakan.

2.2 Flowchart Penelitian

Seperti telah ditulis pada bagian atas bahwa Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif, dengan studi literasi Adapun sumber data dalam penelitian ini adalah data sekunder MPOWER: a dari *The World Health Organization* (WHO), yang sejak tahun 2008 memperkenalkan paket MPOWER: a tersebut [16]. Paket MPOWER: a terdiri dari praktik-praktik terbaik yang digariskan dalam kerangka WHO *Framework Convention on Tobacco Control* (WHO FCTC) untuk membantu mengurangi permintaan tembakau di tingkat negara. Strategi MPOWER: a berisi langkah-langkah berupa M untuk *Monitor tobacco use and prevention policies*, P untuk *Protect*

people from tobacco, O untuk Offer help to quit tobacco use, W untuk Warn about the dangers of tobacco, E untuk Enforce bans on tobacco advertising, promoting and sponsorship, R untuk Raise taxes on tobacco, A untuk Anti-tobacco mass media campaigns. Kumpulan data ini berisi skor nilai variabel M, P, O, W, E, R, dan A dari 195 negara yang merupakan bilangan bulat dengan rentang nilai 1-5 [5]. Selanjutnya juga akan dijabarkan lebih dalam beberapa variabel yang digunakan. Variabel yang digunakan pada penelitian ini, sebagai berikut.

Pada Tabel 1 terdapat 7 variabel yang digunakan. Penilaian penggunaan tembakau dan dampaknya harus diperkuat, pada Variabel M (Monitor tobacco use and prevention policies) dilakukan dengan melakukan survei yang mencakup data tentang penggunaan tembakau secara berkala dengan menggunakan pertanyaan, pengambilan sampel, analisis data, dan teknik pelaporan. Pemantauan seperti ini sangat penting karena dapat memberikan informasi tentang tingkat epidemi tembakau di suatu negara dan bagaimana menyesuaikan kebijakan untuk negara tertentu sesuai kebutuhan sehingga dapat membalikkan epidemi tembakau.

Tabel 1. Deskripsi Variabel

No.	Variabel	Tipe
1	Monitor tobacco use and prevention policies (M)	Numerikal
2	Protect people from tobacco (P)	Numerikal
3	Offer help to quit tobacco use (O)	Numerikal
4	Warn about the dangers of tobacco (W)	Numerikal
5	Enforce bans on tobacco advertising, promoting, and sponsorship (E)	Numerikal
6	Raise taxes on tobacco (R)	Numerikal

7	Anti-tobacco mass media campaigns (A)	Numerikal
---	---------------------------------------	-----------

Variabel P (Protect people from tobacco) dilakukan dengan memberlakukan lingkungan atau tempat bebas rokok untuk melindungi kesehatan orang dari efek berbahaya asap rokok. Perundang-undangan yang dibuat negara terkait tempat bebas rokok di tingkat nasional atau subnasional diberlakukan pada fasilitas perawatan kesehatan, fasilitas pendidikan selain universitas, universitas, fasilitas pemerintah, kantor indoor, restoran, pub dan bar, transportasi umum. Tempat bebas asap rokok sangat penting untuk melindungi non-perokok dan juga dapat mendorong perokok untuk berhenti.

Variabel O (Offer help to quit tobacco use) dilakukan dengan memberlakukan ketersediaan bantuan penghentian tembakau untuk membantu pengguna tembakau berhenti. Negara harus membuat program penghentian tembakau dengan biaya rendah, intervensi yang efektif untuk pengguna tembakau yang ingin lepas dari kecanduannya. Layanan berhenti merokok biasanya tersedia di salah satu tempat berikut: klinik kesehatan atau fasilitas perawatan primer lainnya, rumah sakit, kantor professional kesehatan, komunitas.

Variabel W (Warn about the dangers of tobacco) dilakukan dengan menunjukkan ukuran peringatan tentang bahaya tembakau pada kemasan produk tembakau agar pengguna tembakau di seluruh dunia memahami sepenuhnya risiko terhadap kesehatan mereka. Jadi pada variabel ini mengukur status pelaksanaan perundang-undangan negara yang mengamanatkan peringatan kesehatan pada kemasan produk tembakau berdasarkan kriteria tertentu agar pengguna tembakau mengerti sepenuhnya tingkat risiko kesehatan mereka.

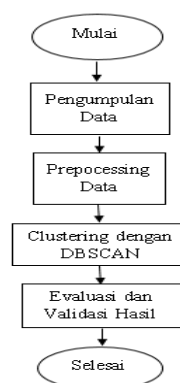
Variabel E (Enforce bans on tobacco advertising, promoting and sponsorship) dilakukan dengan menerapkan larangan iklan, promosi, dan sponsor tembakau sesuai dengan prinsip konstitusional masing-masing negara. Larangan total seperti ini dapat mengurangi konsumsi tembakau dan melindungi orang, khususnya anak-anak muda, dari taktik pemasaran industri.

Menaikkan pajak dan harga adalah cara yang paling efektif untuk mengurangi penggunaan tembakau khususnya untuk kaum muda dan membantu meyakinkan pengguna tembakau untuk berhenti. Variabel R (Raise taxes on tobacco) dilakukan dengan memberikan informasi negara tentang pajak dan harga dinilai untuk menggambarkan tingkat perbandingan pajak atas produk tembakau di negara lain. Pajak yang dikenakan meliputi pajak cukai, pajak pertambahan nilai (“PPN”), bea masuk (ketika rokok diimpor) dan pajak lainnya yang dikenakan. Pajak yang ditetapkan dapat digunakan negara untuk menegakkan kebijakan pengendalian tembakau dan membayar program kesehatan dan sosial masyarakat lainnya.

Variabel A (Anti-tobacco mass media campaigns) telah menjadi komponen kunci dari program pengendalian tembakau di negara-negara. Dengan karakteristik kampanye anti tembakau yang harus dilakukan adalah kampanye yang dilakukan merupakan bagian dari program pengendalian tembakau. Sebelum kampanye, penelitian terkait target audiens harus dilakukan atau ditinjau untuk mendapatkan pemahaman yang menyeluruh. Materi komunikasi kampanye juga telah diuji coba dengan target audiens dan disempurnakan sesuai dengan tujuan kampanye. Untuk waktu tayang (radio, televisi) dan/atau penempatan (papan reklame, iklan cetak, dll.) dilakukan dengan perencanaan yang

menyeluruh untuk menjangkau audiens target secara efektif dan efisien. Lembaga pelaksana dapat bekerja sama dengan wartawan untuk mendapatkan publisitas atau liputan berita untuk kampanye. Evaluasi proses dilakukan untuk menilai seberapa efektif kampanye telah dilaksanakan untuk menilai dampak kampanye.

Berdasarkan uraian di atas, tahapan proses penelitian dapat digambarkan secara ringkas mengikuti alur flowchart sebagai berikut.



Gambar 2. Flowchart Proses Penelitian

3. HASIL DAN PEMBAHASAN

Data yang dikumpulkan untuk penelitian ini adalah data penilaian MPOWER: a pada 195 negara seperti yang telah disebutkan pada bagian sebelumnya. Selanjutnya tahap preprocessing dilakukan seleksi data dan pembersihan data. Seleksi data dilakukan dengan cara memilih data penilaian MPOWER: a pada 195 negara tahun 2018. Sedangkan pembersihan data tidak perlu dilakukan karena pada dataset yang digunakan pada penelitian ini lengkap dan data yang dibutuhkan juga sudah cukup tersedia untuk dapat diolah pada proses clustering. Berikut adalah data yang digunakan pada penelitian ini.

Tabel 2. Penilaian MPOWER:a

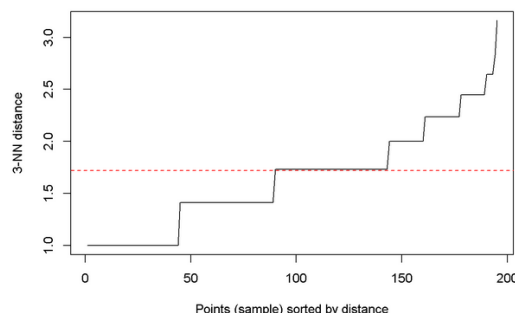
Country	M	P	O	W	E	R	A
Angola	1	3	4	2	2	2	1
Burgundi	1	5	2	3	4	3	2

Benin	3	5	2	3	5	2	2
Burkina Faso	2	5	3	5	4	3	2
Botswana	2	2	4	2	4	3	3
Cental African Public	1	2	2	2	2	3	2
Cote d'Ivoire	1	2	4	2	2	3	4

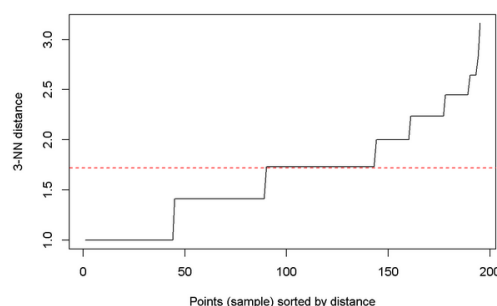
Berdasarkan Gambar 3 terlihat bahwa terdapat 7 variabel yang akan diteliti. Dengan melihat penilaian MPOWER: a tiap negara sebagai contoh Angola untuk variabel M mendapat skor 1, untuk variabel P mendapat skor 3, untuk variabel O mendapat skor 4, untuk variabel W mendapat skor 2, untuk variabel E mendapat skor 2, untuk variabel R mendapat skor 2, dan untuk variabel A mendapat skor 1. Ini mengindikasikan bahwa negara Angola buruk dalam pengelolaan penggunaan tembakau dan pengadaan kampanye anti-tembakau; cukup baik dalam pengelolaan larangan iklan, promosi, dan sponsor tembakau, pajak atas produk tembakau, dan ukuran peringatan tentang bahaya tembakau pada kemasan produk tembakau; baik dalam pengelolaan pemberlakuan lingkungan atau tempat bebas rokok; dan sangat baik dalam penyediaan bantuan penghentian tembakau.

Untuk melakukan clustering DBSCAN adalah dengan memilih parameter MinPts dan Epsilon (ϵ). Parameter MinPts dan Epsilon (ϵ) yang optimum dapat dipilih dengan fungsi kNNdistplot pada Rstudio yang berasal dari packages dbscan. Ide utama dari fungsi kNNdistplot adalah dengan menghitung jarak rata-rata untuk setiap data ke k tetangga terdekatnya (Nearest Neighbors). Pada tahap ini dilakukan untuk mendapatkan nilai K-dist pada nilai k yang bervariasi. K-dist pada penelitian ini akan dicari pada titik k = 3, k = 4, dan

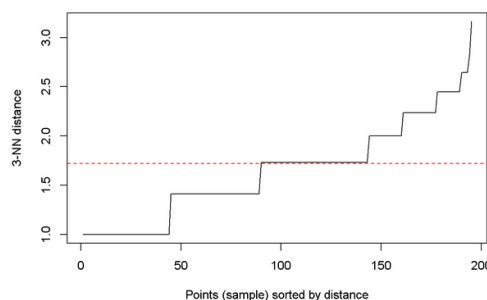
k = 5. Rata-rata jarak yang sudah dapat divisualisasikan dalam plot secara ascending untuk mendapatkan “knee” yang menunjukkan nilai optimal dari Epsilon (ϵ) berdasarkan K yang ditentukan.



Gambar 3. K-dist dengan k = 3



Gambar 4. K-dist dengan k = 4



Gambar 5. K-dist dengan k = 5

Berdasarkan plot diatas dengan menggunakan k = 3 didapat jarak yang optimal yaitu sekitar 1.72. Nilai 1.72 didapat dari posisi “knee” yang terbentuk pada plot. Begitu juga dengan k = 4 didapat jarak yang optimal yaitu sekitar 1.72 dan k = 5 didapat jarak yang optimal yaitu sekitar 2.00. Hasil pencarian nilai epsilon (ϵ) yang optimal diatas dapat

digunakan dalam proses clustering yang mana nilai epsilon (ϵ) adalah 1.72 dengan MinPts 3, nilai epsilon (ϵ) adalah 1.72 dengan MinPts 4, dan nilai epsilon (ϵ) adalah 2.00 dengan MinPts 5. Tahap selanjutnya adalah pembuatan cluster menggunakan function DBSCAN dengan parameter yang telah didapat. Metode pengujian hasil clustering dapat dinilai menggunakan salah satu metode validitas cluster. Metode Davies Bouldin Index (DBI) digunakan untuk menguji hasil cluster dimana suatu cluster akan dianggap memiliki hasil clustering yang optimal jika memiliki DBI minimal [17]. Hasil clustering dengan DBSCAN dengan beberapa nilai MinPts dan Epsilon (ϵ) disajikan pada tabel 3.

Tabel 3. Hasil Clustering DBSCAN

MinPts	3	4	5
Epsilon	1,72	1,72	2,00
Noise	55	75	18
Banyak Cluster yang Terbentuk	9	4	2
Nilai Davies-Bouldin Index	1,12039	1,08118	1,288835

Nilai DBI minimal pada proses clustering DBSCAN penelitian ini diperoleh pada nilai epsilon (ϵ) 1.72 dengan MinPts 4 sebesar 1.08118. Oleh karena itu, parameter yang digunakan untuk membentuk clustering pengendalian konsumsi tembakau berdasarkan skor penilaian MPOWER: a dari WHO untuk setiap negara adalah $\epsilon = 1.72$ dengan MinPts = 4.

Hasil clustering dari data tersebut berbantuan Rstudio dengan 195 observasi terdapat 4 cluster dan sebanyak 75 data diidentifikasi sebagai noise. Cluster 3 merupakan cluster dengan anggota cluster terbanyak yang terdiri dari 79 items. Sedangkan cluster 4 merupakan cluster

dengan anggota cluster paling sedikit yang terdiri dari 4 items.

Untuk pengelompokan menjadi 4 cluster akan ditampilkan pada Tabel 3 dibawah ini.

Tabel 4. Keanggotaan Cluster

Cluster	Negara
Noise	Burundi, Benin, Burkina Faso, Côte d'Ivoire, Democratic Republic of the Congo, Congo, Cabo Verde, Algeria, Eritrea, Ethiopia, Gabon, Ghana, Guinea, Gambia, Kenya, Lesotho, Niger, Nigeria, Rwanda, Senegal, Sao Tome and Principe, Chad, Togo, Uganda, Antigua and Barbuda, Bahamas, Bolivia, Barbados, Cuba, Dominica, Jamaica, Saint Lucia, Mexico, Peru, Paraguay, Trinidad and Tobago, Venezuela, Afghanistan, United Arab Emirates, Bahrain, Djibouti, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Occupied Palestinian Territory, Qatar, Saudia Arabia, Syrian Arab Republic, Yemen, Albania, Andorra, Armenia, Azerbaijan, Iceland, Israel, Monaco, Montenegro, San Marino, Serbia, Tajikista, Bhutan, Brunei Darussalam, China, Micronesia, Japan, Cambodia, Lao People's Democratic Republic, Niue, Nauru, Paulu, Papua New Guinea, Vanuatu
Cluster 1	Botswana, Mali, Mozambique, Mauritania, Eswatini, United Republic of Tanzania, South Africa, Oman, Sudan, Tunisia,

	Bosnia and Herzegovina, Uzbekistan, Maldives, Kiribati, Tuvalu
Cluster 2	Angola, Central African Republic, Comoros, Guinea-Bissau, Equatorial Guinea, Liberia, Malawi, Sierra Leone, South Sudan, Zambia, Zimbabwe, Belize, Dominican Republic, Grenada, Guatemala, Haiti, Saint Kitts and Nevis, Nicaragua, Saint Vincent and The Grenadines, Somalia, Democratic People's Republic of Korea, Marshall Islands
Cluster 3	Cameroon, Madagascar, Mauritius, Namibia, Seychelles, Argentina, Brazil, Canada, Chile, Colombia, Costa Rica, Ecuador, Guyana, Honduras, Panama, El Savador, Suriname, Uruguay, Egypt, Iran, Pakistan, Austria, Belgium, Bulgaria, Belarus, Cyprus, Czechia, Germany, Denmark, Spain, Estonia, Finland, France, United Kingdom of Great Britain and Northern Ireland, Georgia, Greece, Croatia, Hungary, Ireland, Italy, Kazakhstan, Kyrgyzstan, Lithuania, Luxembroug, Latvia, Republic of Moldova, The former Yugoslav Republic of Macedonia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Sweden, Turkmenistan,

	Türkiye, Ukraine, Bangladesh, India, Sri Lanka, Myanmar, Nepal, Thailand, Timor-Leste, Australia, Cook Islands, Fiji, Mongolia, Malaysia, New Zealand, Philippines, Singapore, Solomon Islands, Tonga, Vietnam, Samoa
Cluster 4	United States of America, Switzerland, Indonesia, Republic of Korea

Berdasarkan hasil diatas, maka dapat dilakukan profilisasi tiap kelompok yang terbentuk berdasarkan rata-rata masing-masing variabel pada tiap cluster sebagai berikut.

Tabel 5. Profilisasi Clustering DBSCAN

Cluster	M	P	O	W	E	R	A
1	2, 27	2,2	3, 6	2, 73	4, 2	3, 4	2,1 33
2	1, 68	2,7 3	2, 95	2, 14	2	2, 5	1,7 7
3	3, 63	3,8 62	4, 14	4, 92	4, 22	4, 13	3,3 2
4	4	2,2 5	4, 5	4	2	3, 75	5

- a Cluster 1 : Negara Botswana, Mali, Mozambique, Mauritania, Eswatini, United Republic of Tanzania, dan lain-lain memiliki pengelolaan memberlakukan lingkungan atau tempat bebas rokok yang buruk; penyediaan bantuan penghentian tembakau, pengelolaan ukuran peringatan bahaya tembakau pada kemasan produk tembakau, pengelolaan pajak atas produk tembakau, pengadaan kampanye anti-tembakau yang cukup baik; pengelolaan penggunaan tembakau dan pengelolaan larangan iklan, promosi, sponsor tembakau yang baik.

- b Cluster 2 : Negara Angola, Central African Republic, Comoros, Guinea-Bissau, Equatorial Guinea, Liberia, dan lain-lain memiliki pengelolaan penggunaan tembakau, penyediaan bantuan penghentian tembakau, pengelolaan ukuran peringatan bahaya tembakau pada kemasan produk tembakau, pengelolaan larangan iklan, promosi, dan sponsor tembakau, pengelolaan pajak atas produk tembakau, pengadaan kampanye anti-tembakau yang buruk; dan pengelolaan memberlakukan lingkungan atau tempat bebas rokok yang baik.
- c Cluster 3 : Negara Cameroon, Madagascar, Mauritius, Namibia, Seychelles, Argentina, Brazil, Canada, dan lain-lain memiliki pengelolaan penggunaan tembakau yang cukup baik bagi masyarakat; penyediaan bantuan penghentian tembakau dan pengadaan kampanye anti-tembakau yang baik; pengelolaan memberlakukan lingkungan atau tempat bebas rokok, pengelolaan ukuran peringatan bahaya tembakau pada kemasan produk tembakau, pengelolaan pajak atas produk tembakau, pengelolaan larangan iklan, promosi, dan sponsor tembakau yang sangat baik.
- d Cluster 4 : Negara United States of America, Switzerland, Indonesia, dan Republic of Korea memiliki pengelolaan larangan iklan, promosi, dan sponsor tembakau yang buruk; pengelolaan memberlakukan lingkungan atau tempat bebas rokok yang cukup baik; pengelolaan ukuran peringatan bahaya tembakau pada kemasan produk tembakau, pengelolaan pajak atas produk tembakau yang baik; pengelolaan penggunaan tembakau, penyediaan bantuan penghentian tembakau,

pengadaan kampanye anti-tembakau yang sangat baik.

4. KESIMPULAN

Hasil penelitian adalah pengelompokan pengontrolan penggunaan tembakau berdasarkan skor penilaian MPOWER: a dari WHO untuk setiap negara menggunakan algoritma Density Based Spatial Clustering of Applications with Noise (DBScan). Berdasarkan k-dist graph parameter epsilon (ϵ) dan MinPts yang optimal adalah epsilon (ϵ) sebesar 1.72 dengan MinPts sebesar 4 menghasilkan 4 cluster, 75 data sebagai noise karena memiliki nilai Davies Bouldin-Index terkecil sebagai validitas cluster sebesar 1.08118 jika dibandingkan dengan $\epsilon = 1.72$ MinPts = 3 dan $\epsilon = 2.00$ MinPts = 5. Indonesia sendiri masuk dalam cluster 4 yaitu negara yang memiliki pengelolaan larangan iklan, promosi, dan sponsor tembakau yang buruk; pengelolaan memberlakukan lingkungan atau tempat bebas rokok yang cukup baik; pengelolaan ukuran peringatan bahaya tembakau pada kemasan produk tembakau, pengelolaan pajak atas produk tembakau yang baik; pengelolaan penggunaan tembakau, penyediaan bantuan penghentian tembakau, pengadaan kampanye anti-tembakau yang sangat baik.

DAFTAR PUSTAKA

- [1.] Kurmus, H., & Mohajerani, A. (2020). The toxicity and valorization options of cigarette butts. *Waste Management*, 104, 104-118.
- [2.] Benowitz, N. L. (2009). Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. *Annual review of pharmacology and toxicology*, 49, 57.
- [3.] Badan Narkotika Nasional. (2013, Mei 6). Rokok Gerbang Narkoba. Dipetik Juli 16, 2022, dari Badan Narkotika Nasional: <http://www.bnn.go.id/read/artikel/10852/rokok-gerbang-narkoba>.

- [4.] Shimatani, K., Ito, H., Matsuo, K., Tajima, K., & Takezaki, T. (2020). Cumulative cigarette tar exposure and lung cancer risk among Japanese smokers. *Japanese Journal of Clinical Oncology*, 50(9), 1009-1017.
- [5.] World Health Organization [internet]. Tobacco. Geneva: World Health Organization; 2017.
- [6.] Valarmathy, N., & Krishnaveni, S. (2020). A novel method to enhance the performance evaluation of DBSCAN clustering algorithm using different distinguished metrics. *Materials Today: Proceedings*.
- [7.] Sheridan, K., Puranik, T. G., Mangortey, E., Pinon-Fischer, O. J., Kirby, M., & Mavris, D. N. (2020). An application of dbSCAN clustering for flight anomaly detection during the approach phase. In *AIAA Scitech 2020 Forum* (p. 1851).
- [8.] Dudik, J. M., Kurosu, A., Coyle, J. L., & Sejdić, E. (2015). A comparative analysis of DBSCAN, K-means, and quadratic variation algorithms for automatic identification of swallows from swallowing accelerometry signals. *Computers in biology and medicine*, 59, 10-18.
- [9.] Sitompul, B. J. D., Sitompul, O. S., & Sihombing, P. (2019, June). Enhancement clustering evaluation result of davies-bouldin index with determining initial centroid of k-means algorithm. In *Journal of Physics: Conference Series* (Vol. 1235, No. 1, p. 012015). IOP Publishing.
- [10.] Wijaya, Y. A., Kurniady, D. A., Setyanto, E., Tarihoran, W. S., Rusmana, D., & Rahim, R. (2021). Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities. *TEM J*, 10(3), 1099-1103.
- [11.] Francis, Z., Villagrasa, C., & Clairand, I. (2011). Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. *Computer methods and programs in biomedicine*, 101(3), 265-270.
- [12.] Sawant, K. (2014). Adaptive methods for determining DBSCAN parameters. *International Journal of Innovative Science, Engineering & Technology*, 1(4), 329-334.
- [13.] Viswanath, P., & Babu, V. S. (2009). Rough-DBSCAN: A fast hybrid density based clustering method for large data sets. *Pattern Recognition Letters*, 30(16), 1477-1488.
- [14.] Utomo, W. (2021). The comparison of k-means and k-medoids algorithms for clustering the spread of the covid-19 outbreak in Indonesia. *ILKOM Jurnal Ilmiah*, 13(1), 31-35.
- [15.] Lomo, S. I., Darmawan, E., & Sugiyarto, S. (2021). Cluster analysis of type II diabetes mellitus patients with the fuzzy c-means method. *Annals of Mathematical Modeling*, 13(1), 32-39.
- [16.] Hoang, V. M., Tran, T. N., Vu, Q. M., Nguyen, T. T. M., Le, H. C., Vu, D. K., ... & Kim, B. G. (2016). Tobacco control policies in Vietnam: review on MPOWER implementation progress and challenges. *Asian Pacific Journal of Cancer Prevention*, 17(sup1), 1-9.
- [17.] Hoang, V. M., Tran, T. N., Vu, Q. M., Nguyen, T. T. M., Le, H. C., Vu, D. K., ... & Kim, B. G. (2016). Tobacco control policies in Vietnam: review on MPOWER implementation progress and challenges. *Asian Pacific Journal of Cancer Prevention*, 17(sup1), 1-9.